

Private

BEST AVAILABLE COPY

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2006, 01:28:04 ; Search time 316 Seconds
(without alignments)
3061.919 Million cell updates/sec

Title: US-10-771-931-1
Perfect score: 5357
Sequence: 1 MTNCKYKRLKSLVGLVSVGT.....PTIIGAAGLSKRRDTRGN 1046

Scoring table:

BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: UniProt 7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5357	100.0	1046	2	084941_STRPY
2	3575.5	66.7	1027	2	093T51_STRPY
3	3501	65.4	1026	2	048R34_STRPM
4	3442.5	63.9	1023	2	093T53_STRPY
5	3371.5	62.9	1025	2	054507_STRPY
6	3370	62.7	1026	2	09X3R6_STRPY
7	3357.5	62.7	1025	2	P72532_STRPY
8	3164.5	59.1	1013	2	091953_STRPY
9	3132.5	58.5	1015	2	093T50_STRPY
10	2973.5	55.5	1013	2	093T52_STRPY
11	2878.5	52.9	954	2	06YB06_STRPY
12	2834	53.7	1022	2	093T54_STRPY
13	2779.5	51.9	963	2	09X249_STRPY
14	2654.5	49.6	1029	2	09X242_STRPY
15	2571	48.0	862	2	09R222_STRPY
16	2531	47.2	923	2	09S312_STRPY
17	2488	46.4	865	2	09S449_STRPY
18	2465	46.0	872	2	09S430_STRPY
19	2460.5	45.9	872	2	09S443_STRPY
20	2439.5	45.5	854	2	09S318_STRPY
21	2329	43.5	861	2	09R304_STRPY
22	2138.5	39.9	1091	2	006556_STRPY
23	2102	37.2	746	2	08V935_STRPY
24	1458.5	27.2	409	2	09X9B9_STRPY
25	1391.5	26.0	456	2	09XCK8_STRPY
26	1388.5	25.9	471	2	09R2F8_STRPY
27	1376.5	25.7	447	2	09R007_STRPY
28	1343	25.1	432	2	09XCK2_STRPY
29	1331.5	24.9	439	2	09R008_STRPY
30	1276	23.8	454	2	09S418_STRPY
31	1262	23.6	423	2	09R2F6_STRPY

32	1247.5	23.3	435	2	09X9C0_STRPY
33	1238.5	23.1	437	2	09R2F7_STRPY
34	1231.5	23.0	441	2	09R2F9_STRPY
35	1229	22.9	410	2	09P009_STRPY
36	1219	22.8	253	2	09S492_STRPY
37	1191	22.2	434	2	09X6H2_STRPY
38	1190.5	22.2	430	2	09X613_STRPY
39	1093.5	20.4	938	2	03ZJF8_STRSU
40	1064	19.9	421	2	09R2F5_STRPY
41	1025.5	19.1	383	2	09L5D7_STRPY
42	845.5	15.8	564	2	0300K5_STRSU
43	806.5	15.1	261	2	09R4J5_STRPY
44	774.5	14.5	258	2	09X6H5_STRPY
45	708.5	13.2	250	2	09X6H0_STRPY

ALIGNMENTS

RESULT 1
084941_STRPY PRELIMINARY; PRT; 1046 AA.
AC 084941;
DT 01-NOV-1998, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1998, sequence version 1.
DT 07-FEB-2006, entry version 25.
DE Serum opacity factor.
GN Name=sof2;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=T2MR;
RX MEDLINE=99232508; PubMed=10216862;
RA Courtney H.S., Hasty D.L., Li Y., Chiang H.C., Thacker J.L.,
RA Dale J.B.;
RT "Serum opacity factor is a major fibronectin-binding protein and a
RT virulence determinant of M type 2 Streptococcus pyogenes.";
RL Mol. Microbiol. 32:89-98 (1999).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC
CC EMBL; AF019890; AAC32596.1; -; Genomic DNA.
CC
CC GO; GO:0009986; C:cell surface; IEA.
CC GO; GO:0016020; C:membrane; IEA.
CC GO; GO:0007155; P:cell adhesion; IEA.
CC InterPro; IPR004237; Fibron repeat_bd.
CC InterPro; IPR005877; Gpos YSIRK.
CC InterPro; IPR001899; Gram_pos_anchor.
CC InterPro; IPR002035; VWP_A.
CC Pfam; PF02986; Fn_bind; 3.
CC Pfam; PF00746; Gram_pos_anchor; 1.
CC SMART; SM00327; VWA; 1.
CC TIGRfam; TIGR01168; YSIRK signal; 1.
CC PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
CC PROSITE; PS50234; VWFA; 1.
CC SEQUNCE 1046 AA; 114914 MW; B12CF6B68059B62 CRC64;
SQ
Query Match 100.0%; Score 5357; DB 2; Length 1046;
Best Local Similarity 100.0%; Pred. No. 3.1e-205;
Matches 1046; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTNCKYKRLKSLVGLVSVGTMLIAPLVQVBSASTTSTSTASAGVTSGLASSETS 60
DB 1 MTNCKYKRLKSLVGLVSVGTMLIAPLVQVBSASTTSTSTASAGVTSGLASSETS 60
QY 61 GAATTAATTTATNGPGSTPAVAATPOPOQAIAPVAATSTASASSSGKAPQAVT 120

Db 61 GAATATTATTATNGGPOSTPAVAEATPOQAOIAVAAATSSASSSSDDKAPAVT 120
 QY 121 SSTSSTPAAASNGSNOBASAEPTQMEVEKYTVDKENSKINIKDKPTKTSQSVNNE 180
 Db 121 SSTSSTPAAASNGSNOBASAEPTQMEVEKYTVDKENSKINIKDKPTKTSQSVNNE 180
 QY 181 KOTKLRNBDGKRDIDVTRVTKNEDGTIDVTVVTKPQIDBQDVMAILLVSKKMS 240
 Db 181 KOTKLRNBDGKRDIDVTRVTKNEDGTIDVTVVTKPQIDBQDVMAILLVSKKMS 240
 QY 241 DDFNNAKNTKICLVTKLTSKASNSDNDKHKYNSRNSVRLATEYRISNPIDISGTERO 300
 Db 241 DDFNNAKNTKICLVTKLTSKASNSDNDKHKYNSRNSVRLATEYRISNPIDISGTERO 300
 QY 301 LDKLLDDLRKKAANYDWGVDLOGAIIKABEINKEKEKFGRRHIVLFSQGESYFSYE 360
 Db 301 LDKLLDDLRKKAANYDWGVDLOGAIIKABEINKEKEKFGRRHIVLFSQGESYFSYE 360
 QY 361 LONSVAEDKTKSRLSGAVTSSNPLPWPIFNHTKNTIDMLDVKNLVKLGOTLGIAGL 420
 Db 361 LONSVAEDKTKSRLSGAVTSSNPLPWPIFNHTKNTIDMLDVKNLVKLGOTLGIAGL 420
 QY 421 DMLQSTLSISTSSLAGAFLGGGSLTEYLLKRYSGDKENQPYTKRVEGYHSHF 480
 Db 421 DMLQSTLSISTSSLAGAFLGGGSLTEYLLKRYSGDKENQPYTKRVEGYHSHF 480
 QY 481 SERKKTGEIIPKSEIEPKIELEFNNKNNODKSWTEWIDFKLSLTERIQAKOETLMKLL 540
 Db 481 SERKKTGEIIPKSEIEPKIELEFNNKNNODKSWTEWIDFKLSLTERIQAKOETLMKLL 540
 QY 541 ETLFYREHYNNHNSAIAAQAQOEGITFYSVUTDLKTSKVKQVSTEDKKE 600
 Db 541 ETLFYREHYNNHNSAIAAQAQOEGITFYSVUTDLKTSKVKQVSTEDKKE 600
 QY 601 KOREDEKERNKEDNYLKOMSESGDFEDVDKAEKFOILNTVVTETFEEDGVANON 660
 Db 601 KOREDEKERNKEDNYLKOMSESGDFEDVDKAEKFOILNTVVTETFEEDGVANON 660
 QY 661 SNOVSENNNSLHNSYVTHKASDASWMSLYSNKESLWTISKEQLKEAFKNSSLTF 720
 Db 661 SNOVSENNNSLHNSYVTHKASDASWMSLYSNKESLWTISKEQLKEAFKNSSLTF 720
 QY 721 KTKLVNKKOJLDKNNKRTDSTENKTSVTKOIISNTVYKINNOYKNGLDVVKLT 780
 Db 721 KTKLVNKKOJLDKNNKRTDSTENKTSVTKOIISNTVYKINNOYKNGLDVVKLT 780
 QY 781 YKKEIVPVDVEGEVPIPEKPLVEBMTPLYPALPNVPTPDIPTPOLKDEDEISGGHG 840
 Db 781 YKKEIVPVDVEGEVPIPEKPLVEBMTPLYPALPNVPTPDIPTPOLKDEDEISGGHG 840
 QY 841 PSVDIYEDTGTGABGAONGVSTQENRDIYDITEDTQPGMSGNDATVEEDTAPKRP 900
 Db 841 PSVDIYEDTGTGABGAONGVSTQENRDIYDITEDTQPGMSGNDATVEEDTAPKRP 900
 QY 901 DVLVGGOSPIDITETOPSVSGSNDATVEEDTVKRPDSLVGGOSDPDITEDTQPGM 960
 Db 901 DVLVGGOSPIDITETOPSVSGSNDATVEEDTVKRPDSLVGGOSDPDITEDTQPGM 960
 QY 961 SSGNGATVLEEDTRPKRVFHFDEPOAPEKNEPQSLSLPOAPVYAAHHLPASGDREA 1020
 Db 961 SSGNGATVLEEDTRPKRVFHFDEPOAPEKNEPQSLSLPOAPVYAAHHLPASGDREA 1020
 QY 1021 SFTIAPPTIIGAAGLISKGRDTEGN 1046
 Db 1021 SFTIAPPTIIGAAGLISKGRDTEGN 1046

RESULT 2
 093751 STREPT PRELIMINARY; PRT; 1027 AA.
 093751 STREPT Integrated into UniproctKB/TREMBL.

DT 07-FEB-2006, entry version 16.
 DE Serum opacity factor VT21.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 RN [1]_Taxid=1314;
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21642587; PubMed=11782509;
 RA Gillen C.M., Towers R.J., McMillan D.J., Delvecchio A.,
 RA Sripathi K.S., Currie B., Kreilkemeyer B., Chhatwal G.S.,
 RA Walker M.J.;
 RT Immunological response mounted by Aboriginal Australians living in
 RT the Northern Territory of Australia against Streptococcus pyogenes
 RT serum opacity factor."
 RL Microbiology 148:169-178(2002).
 CC
 CC Copyrighted by the Uniproct Consortium, see <http://www.uniproct.org/terms>
 CC distributed under the Creative Commons Attribution-NonDerivs License
 CC
 CC EMBL; AF367014; AK529681; -; Genomic DNA.
 DR GO; GO:0009986; C:cell surface; IEA.
 DR GO; GO:0009275; C:cell wall (sensu Gram-positive Bacteria); IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR004237; F1dron repeat_bd.
 DR InterPro; IPR005877; Gpos_YSTRK.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR InterPro; IPR002035; VMP_A.
 DR Pfam; PF02986; Fb_bind; 3.
 DR Pfam; PF00746; Gram_pos_anchor; 1.
 DR Pfam; PF00092; VMA; 1.
 DR SMART; SM00327; VMA; 1.
 DR TIGRFAMs; TIGR01168; YSTRK_signal; 1.
 DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
 DR PROSITE; PS0234; VMPA; 1.
 SQ SEQUENCE 1027 AA; 112542 MW; 048239E8598A4C7A CRC64;

Query Match 66.7%; Score 3575.5; DB 2; Length 1027;
 Best Local Similarity 70.1%; Pred. No. 2.8e-134;
 Matches 751; Conservative 98; Mismatches 153; Indels 69; Gaps 26;

QY 1 MNCCYKLRKXSVGVSVGMILAPTVVORVAST---TSENSTASAGVSGTASB 57
 Db 1 MNCCYKLRKXSVGVSVGMILAPTVVORVAST---TSENSTASAGVSGTASB 57
 QY 58 TSSGAATVTTATTT-----ATTNGGPOSTPAVAEATPOQAOIAVAAAT 102
 Db 58 TSSGAATVTTATTT-----ATTNGGPOSTPAVAEATPOQAOIAVAAAT 102
 QY 103 STSSASSSSDGAPOAVTSS7SPSTPAAASNGSNOBASAEPTQMEVEKYTVDKENSK 162
 Db 103 STSSASSSSDGAPOAVTSS7SPSTPAAASNGSNOBASAEPTQMEVEKYTVDKENSK 162
 QY 112 SSS-SSSSDGIIPK--TAITSPTPVAASNN-SNOBASAEPTQMEVEKYTVKEN-Q 166
 Db 112 SSS-SSSSDGIIPK--TAITSPTPVAASNN-SNOBASAEPTQMEVEKYTVKEN-Q 166
 QY 163 LNIKDKPTKSGSSVNNKOTKILNRPQKARDIYDTRVTKNEDGTIDVTVVTKPQI 222
 Db 163 LNIKDKPTKSGSSVNNKOTKILNRPQKARDIYDTRVTKNEDGTIDVTVVTKPQI 222
 QY 167 AKYKGDGDKPKGSSV-DKDKTILNRNDGKORDIYDTRVTKNEDGTIDVTVVTKPQI 225
 Db 167 AKYKGDGDKPKGSSV-DKDKTILNRNDGKORDIYDTRVTKNEDGTIDVTVVTKPQI 225
 QY 223 DEGADVMAILLVSKKMSDDFNNAKNTKICLVTKLTSKASNSDNDKHKYNSRNSVRLMT 282
 Db 223 DEGADVMAILLVSKKMSDDFNNAKNTKICLVTKLTSKASNSDNDKHKYNSRNSVRLMT 282
 QY 226 DEGAADVMAILLVSKKMSDDFNNAKNTKICLVTKLTSKASNSDNDKHKYNSRNSVRLMT 279
 Db 226 DEGAADVMAILLVSKKMSDDFNNAKNTKICLVTKLTSKASNSDNDKHKYNSRNSVRLMT 279
 QY 283 FTRISNPIDISGTEROJLDKLLDRKKAANYDWGVDLOGAIIKABEINKEKEKFG 342
 Db 283 FTRISNPIDISGTEROJLDKLLDRKKAANYDWGVDLOGAIIKABEINKEKEKFG 342
 QY 280 FYREISKYIDISGTEROJLDKLLDRKKAANYDWGVDLOGAIIKABEINKEKEKFG 339
 Db 280 FYREISKYIDISGTEROJLDKLLDRKKAANYDWGVDLOGAIIKABEINKEKEKFG 339
 QY 343 KRRHIVLFSQGESYFSYELONSVAEDKTKSRLSGAVTSSNPLPWPIFNHTKNTIDML 402
 Db 343 KRRHIVLFSQGESYFSYELONSVAEDKTKSRLSGAVTSSNPLPWPIFNHTKNTIDML 402
 QY 340 KRQHTVLSQGESYFSYELONSVAEDKTKSRLSGAVTSSNPLPWPIFNHTKNTIDML 396
 Db 340 KRQHTVLSQGESYFSYELONSVAEDKTKSRLSGAVTSSNPLPWPIFNHTKNTIDML 396
 QY 403 DDKVNLVLDGOTLGIAGLNDQSTLSISTSSLAGAFLGGGSLTEYLLKRYSGDKENQ 462
 Db 403 DDKVNLVLDGOTLGIAGLNDQSTLSISTSSLAGAFLGGGSLTEYLLKRYSGDKENQ 462
 QY 397 DDKVNLVLDGOTLGIAGLNDQSTLSISTSSLAGAFLGGGSLTEYLLKRYSGDKENQ 456
 Db 397 DDKVNLVLDGOTLGIAGLNDQSTLSISTSSLAGAFLGGGSLTEYLLKRYSGDKENQ 456

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd

OM protein - protein search, using sw model

Run on: July 6, 2006, 01:31:29 ; Search time 48 Seconds
(without alignments)
2096.724 Million cell updates/sec

Title: US-10-771-931-1

Sequence: 1 MTNCKYKLRKLSVGLSVGT.....PTIIGAAGLLSKRRDTEGN 1046

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :
1:  p1r_80:*
2:  p1r2:*
3:  p1r3:*
4:  p1r4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3346.5	62.5	1025	2	S69790	fibronectin-bidin
2	2138.5	39.9	1091	2	S33950	fibronectin-binding
3	409.5	7.6	655	2	S40043	adhesin - Streptoc
4	394	7.4	638	2	S54418	fibronectin-bidin
5	354	6.6	961	2	G90053	fibronectin-bidin
6	351.5	6.6	1018	2	A32192	fibronectin-binding
7	339	6.3	1038	2	H90053	fibronectin-binding
8	338.5	6.3	940	2	S19702	fibronectin-bidin
9	325.5	6.1	1039	2	T30856	protecin P2 - Strep
10	306	5.7	1117	2	S33851	fibronectin-bidin
11	302	5.6	1164	1	PCSOXG	IGA Fc receptor pr
12	289	5.4	1134	2	AG0234	fibronectin-bidin
13	282	5.3	1092	2	S42798	fibronectin-bidin
14	268.5	5.0	1875	2	G38173	myosin-like protei
15	263.5	4.9	1726	1	SAZQCM	major merozoite su
16	263.5	4.9	1726	2	A45948	major merozoite su
17	263	4.9	2136	2	T18296	myosin heavy chain
18	257.5	4.8	2481	1	D90011	Fmbp protein [limp
19	248.5	4.6	2722	2	T20552	hypothetical prote
20	248	4.6	2663	1	S58261	centromere protein
21	247	4.6	1526	2	A45605	mature-parasite-in
22	245	4.6	1086	2	S16752	major merozoite su
23	241.5	4.5	853	2	T51505	hypothetical prote
24	241	4.5	2269	2	T28677	hypothetical prote
25	240.5	4.5	1306	2	S22624	aggregation protei
26	240	4.5	1127	2	T28317	ORF MSV156 hypothe
27	230	4.5	1141	2	E89824	hypothetical prote
28	239.5	4.5	1125	2	E90598	membrane nuclease,
29	239.5	4.5	2508	2	S61441	surface-associated

30	239.5	4.5	2738	2	E86320	protein P07A1.6 (
31	238.5	4.5	1163	2	G97236	Atase involved in
32	237	4.4	1192	2	A71625	probable secreted
33	237	4.4	1396	1	NH501F	aggregation protei
34	236.5	4.4	1166	2	T26860	fibronogen-binding
35	236.5	4.4	3488	2	T34418	hypothetical prote
36	236	4.4	2532	2	T18477	hypothetical prote
37	235	4.4	1381	1	S45781	probable calcium-b
38	235	4.4	1822	2	S33441	LF protein - Strept
39	234	4.4	2401	2	T28676	rophory protein -
40	233.5	4.4	816	2	S64439	hypothetical prote
41	232.5	4.3	7160	2	T27935	hypothetical prote
42	231.5	4.3	1690	2	T11030	microtubule bindin
43	231	4.3	1256	1	A34829	miramidase-release
44	231	4.3	1473	2	A5186	salivary agglutinin
45	230.5	4.3	1701	2	A26868	major merozoite su

ALIGNMENTS

RESULT 1

fibronectin-binding protein II precursor - Streptococcus pyogenes
 C:Species: Streptococcus pyogenes
 C>Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S69790; S72666
 R:Kreikemeyer, B.; Talay, S.R.; Chhatwal, G.S.
 Mol. Microbiol. 17, 137-145, 1995
 A:Title: Characterization of a novel fibronectin-binding surface protein, in group A streptococci
 A:Reference number: S69790; MUID:96020668; PMID:7476200
 A:Accession: S69790
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-1025 <RKB>
 A:Cross-references: UNIPROT:P72532; UNIPARC:UPI000017ACTA; EMBL:X83303; NID:g1070387; F:1993-1016/Domain: membrane anchor cleavage motif
 A:Experimental source: strain A75
 A>Note: the authors translated the initiation codon TTG for residue 1 as Leu
 R:Kreikemeyer, B.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: S72666
 A:Accession: S72666
 A:Molecule type: DNA
 A:Residues: 1-19, 'T', 21-145, 'T', 147-1025 <KRM>
 A:Cross-references: UNIPARC:UPI000004F5EB; EMBL:X83303; NID:g1070387; PIDN:CAA58282.1;
 C:Genetics:
 A:Gene: sfbII
 A:Start codon: TTG
 C:Keywords: fibronectin binding; membrane bound
 R:1-34/Domain: signal sequence #status predicted <SIG>
 R:35-1025/Product: fibronectin-binding protein II #status predicted <MAT>
 F:848-868/Region: fibronectin binding
 F:869-907/Region: fibronectin binding
 F:908-946/Region: fibronectin binding
 F:950-994/Region: membrane anchor cleavage motif
 F:995-1016/Domain: transmembrane #status predicted <TM>

	Query Match	62.5%	Score 3346.5	DB 2	Length 1025
	Best Local Similarity	64.0%	Pred. No.3,1e-130		
	Matches	682	Conservative 142	Mismatches 181	Indels 61
				Gaps	19
QY	1	MTNCKYKLRKIISVGLVSVGNMLIAPTVLVGEVSAST--TSTETSTASAG-----VTSG	52		
Db	1	MTNCKYKLRKIISVGLVSVGNMILAPTVLGRVMSASTETSSASTSTSTAETSPYTGSG	60		
QY	53	TAASTGTGAATTAITTTATTNGPOSTPAVARATPPQAQIAPIVAAA---TSTSASSS	110		
Db	61	TAAAS-GASGKA-----TVATANGPOSAPAFTSETPPQAQAAPALSAFTTVISSSSSD	113		
QY	111	SDGKAPAVTSTPSPTPALAASNSGNOEASAEETEPQTMEYEKTVPDKENKLINIKDXT	170		
Db	114	SDAKTPKA-AATTSSATVAASPNSGNNKEANAEPAPQMMDVEQYRKIDENSITYAD---	169		

QY 171 PKTGSVNNKEDTKLIRNBDKLDIVDVTYRTVKTNEEDGTIDVTYVTKPKQIDEGADVMA 230
 DB 170 -----KAKQKIRBDBNDPKDLPDVKREKVDNDGDLDTLTKMPQIDEGADVMA 221
 QY 231 LLDVSKKMSDEDFNNANKIKLVTLTSSKASNSNDDEKYNSSRNLMTFYEKREISNP 290
 DB 222 LLDVSKKMSDEDFNNANKIKLVTLTSSKASNSNDDEKYNSSRNLMTFYEKREISNP 290
 QY 291 IDISGTEBOLDKLDLDRKAKANVWDVLOGAHKAREIFNKEKKGKGRHVLV 350
 DB 278 TEL---TTKQVDAKKEWDAKKQMDWGLQAGAHKAREIFNKEKKGKGRHVLV 332
 QY 351 SGGESTSYLQNSVREDTKLSTLSGAVTSSNPLPMPPIFNTHKNIMLDDVKNLVK 410
 DB 333 SGGESTSYLQNSVREDTKLSTLSGAVTSSNPLPMPPIFNTHKNIMLDDVKNLVK 410
 QY 411 LGOTLGIAGLDNOSTLSLSTGSSLAGAFGGSLTEYLTLKEKYSGDLEKNOFDYTKR 470
 DB 392 WEGTLGIBGLDNLNTLKLGAAGAGIVGFLGGSLTEYLTLKEKYSGDLEKNOFDYTKR 451
 QY 471 VGBGYHPSFPERKKTGEIPKSEIEPKIKELFENNNKNDKSWTEWIFDKLSTLTKRQ 530
 DB 452 VGBGYHPSFPERKKTGEIPKSEIEPKIKELFENNNKNDKSWTEWIFDKLSTLTKRQ 530
 QY 531 AKQETLMKLELYLTKGEHYHYNNLSAIAKAAQOGITFYSVDVDTLTKTSKRVKQ 590
 DB 507 AKQETLMKLELYLTKGEHYHYNNLSAIAKAAQOGITFYSVDVDTLTKTSKRVKQ 590
 QY 591 VESBEDKKKORHDEKRNKEKPNYKQSEGGKDFPDVDKAKKEKDLITNVTYET 650
 DB 567 VESBEDKKKORHDEKRNKEKPNYKQSEGGKDFPDVDKAKKEKDLITNVTYET 650
 QY 651 FEDGVNVKNSMVOYSSENNNSLHNSYKVTAKAASDASWMSLY-SNKSLLTWTSKQK 709
 DB 627 FGNVSVSSGSKMSLGSNSG--SSSRKYSKGRDSGLSLSTGSKESLWTSKQK 684
 QY 710 BAFKNSSLTPKTKLQVKNKQKLDKNKR-----TKDSTENKTSYTKOISNTVYK 763
 DB 665 QAFEGKELTLYLKLKDKDKRETKLKQOESRRIRKKAASSESNTVETLISKISYK 744
 QY 764 INN-QEYVGNLDDVKTLYTKEFVVPDVEGEVY-PIPEKPLVEMTPLYPAIDNYP 820
 DB 745 INNGDTONNNKLEBVSYSKFMPIPELDIEVVVPQPEKPLVEMTPLYPAIDNYP 801
 QY 821 DIPQLPREDLEISGHPGVDIVEDTGTABEGAGNVSTOENRDPYDITIEDTOP 880
 DB 802 --PTQLPREDELEISGHPGVDIVEDTGTABEGAGNVSTOENRDPYDITIEDTOP 880
 QY 881 GMSGNDATVVEEDTAPKRPDVLVGGQSPDITIEDTOPSVSGSNDATVVEEDTAPKRPD 940
 DB 860 GMSGNDATVVEEDTAPKRPDVLVGGQSPDITIEDTOPSVSGSNDATVVEEDTAPKRPD 940
 QY 941 SLVGGQSDIDITIEDTOPGMSGNGATVIEBTPRPKRVFHPDNEPQAPKPEKNEPBLSP 1000
 DB 920 ILVGGQSDIDITIEDTOPGMSGNGATVIEBTPRPKRVFHPDNEPQAPKPEKNEPBLSP 1000
 QY 1001 QAPVYKAHLTPASGDKREASPTIATPITIGAGLISKRRTDTEG 1046
 DB 980 QAPVYKAHLTPASGDKREASPTIATPITIGAGLISKRRTDTEG 1046

RESULT 2

S33850
 fibronectin-binding protein - Streptococcus dysgalactiae
 C:Species: Streptococcus dysgalactiae
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
 C:Accession: S33850, S32632
 R:Linkgren, P.E.; McGavin, M.U.; Sigaard, C.; Guse, B.; Gursiddappa, S.; Hoesek, M.; Li
 Eur. J. Biochem. 214, 819-827, 1993
 A:Title: Two different genes coding for fibronectin-binding proteins from Streptococcus
 A:Reference number: S33850, MUD:93307299, PMID:8319691

A:Molecule type: DNA
 A:Residues: 1-1091 <LIN>
 A:Cross-references: UNIPROT:Q06556; UNIPARC:UPI00000B92C0; EMBL:Z22150; NID:9288966; PI
 Query Match 39.9%; Score 2138.5; DB 2; Length 1091;
 Best Local Similarity 42.3%; Pred. No. 1.5e-80;
 Matches 487; Conservative 184; Mismatches 310; Indels 171; Gaps 28;

QY 1 MTNCKYKRLKSLGVGVSTGMLAPVTVQVSASTSTSTSTASAGVGTGTAASSTGS 60
 DB 1 MTNCKYKRLKSLGVGVSTGMLAPVTVQVSASTSTSTSTASAGVGTGTAASSTGS 60
 QY 61 GAAVTAATTTATTCGPGSTPAVVAATPOPAQIPVAAATSTSSASSSDG----- 113
 DB 37 -ASVTTESPAIQTEHDQSGAEALIEPTAPQTPSPVSAVPEKAAAMDEKIAEAPAH 95
 QY 114 -KAPQ--VTSSTSPSTPAASNS-----GNOEASAEETEPQMEVEKYTVKEN 160
 DB 96 EPAPKASVQAEASPPKAEATNTGQPTVTEQARSRSKAAIAIQTIEVEKLEVDKEN 155
 QY 161 SKNIDKQKTPKTSVNNKEDTKLIRNBDKLDIVDVTYRTVKTNEEDGTIDVTYVTKPK 220
 DB 156 SSLTVNDG-----EKDKQKILKRDGNQDIPDISRDYKVNQDGTMDVTLVYKPK 204
 QY 221 QIDEGADVMA LLDVSKKMSDEDFNNANKIKLVTLTSSKASNSNDDEKYNSSRNLMT 280
 DB 205 QIDEGADVMA LLDVSKKMSDEDFNNANKIKLVTLTSSKASNSNDDEKYNSSRNLMT 280
 QY 281 MTTPRAISNIDISGTEBOLDKLDLDRKAKANVWDVLOGAHKAREIFNKEKKGKGRHVLV 339
 DB 264 IDPFRKAGSSTDLISGMDAKKIDERNKNEVWKAADIDNGVVDQAGAHKAREIFNKEKKGKGRHVLV 339
 QY 340 KFGKRRHIVLFSGQESTSYLQNSVREDTKLSTLSGAVTSSNPLPMPPIFNTHKNIML 399
 DB 324 RSGKRGHIVLFSGQESTSYLQNSVREDTKLSTLSGAVTSSNPLPMPPIFNTHKNIML 399
 QY 400 DMLDDVKNLVKLGOTLGIAGLDNOSTLSLSTGSSLAGAFGGSLTEYLTLKEKYSGDLE 459
 DB 381 NVVNDAKKLDIPFLNKLGISQFNGAVNDVAVGNTLLGLGSPFGIKMPLDVLISLADLETSK 440
 QY 460 LKENQPDYTKRVGSGHPSFPERKKTGEIPKSEIEPKIKELFENNNKNDKSWTEWIF 519
 DB 441 INSKFPYSRSGVGVNFRSTFDR-EVDKVGFKILVETK---GNLAKKQPKQTDVRL- 495
 QY 520 DKLSTLTKRQAKQETLMKLELYLTKGEHYHYNNLSAIAKAAQOGITFYSVDVDTLTKTSKRVKQ 590
 DB 496 SSLGINSIKELIQOMMDKLDNLNFRQYQFVYNNLSAQAARMAABEGIKFYAVDYTE 555
 QY 580 LKTSKRVKQVES-----TEDKKKORHDEKRNKEKPNYKQSEGGKDFPDVDKAKKEKDLITNVTYET 650
 DB 556 PERIAKEINSQYSEAYTNHLKKAABEAREBAKGRNEKFDKYLKEMSSQK-FEKDVEDP 614
 QY 636 EKPQDILNNVTVTEFEDGVNVKNSMVOYSSENNNSLHNSYKVTAKAASDASWMSLY-SNKSLLTWTSKQK 709
 DB 615 EKPQDILNNVTVTEFEDGVNVKNSMVOYSSENNNSLHNSYKVTAKAASDASWMSLY-SNKSLLTWTSKQK 709
 QY 694 --SNKSLLTWTSKQKAKQETLMKLELYLTKGEHYHYNNLSAIAKAAQOGITFYSVDVDTLTKTSKRVKQ 590
 DB 665 SSTTMSSTLWLSKDKKALQSGETLTLEYKLIKHKQFKLAPQTRSKESLDTSENKKS 724
 QY 751 VTKDIISNTVANKINQVYKANKLDVGLTYTKEFVVPDVEGEVYPIPEKPLVEMTPLYPAIDNYP 820
 DB 725 VTEKIVTSVVKIKINDKEVKEKLDVSLVSKETVRKQVEPNV--PDTQEKELPDL 781
 QY 811 VYALINYP-----TPDIP-----TPQLPREDELEISGHPGVDIVEDTGTABEGAGNVSTOENRDPYDITIEDTOP 880
 DB 782 ADSBSQSPISPTPLPSPSPVSTPSTBPGFEGNNL-GGGSBRLITIEDQSGMSGQN 840
 QY 856 ---GAQNGVVSQENRDPV-----DITIEDTOPGMSGNDATVVEEDTAPKRPDVLV 1000
 DB 841 PDSGNETVVEDTQTEQEDVLDGAPQGVLDITFEDSOPGMSGNSHITIEDSKPSQEDVIT 1000
 QY 905 GQGSPPDITIEDTOPSVSGSNDATVVEEDTAPKRPDVLV 1000

171 PRTGSSVNNKEDTKLIRNRDGLADIVDTRTVKTNEBDITDVTVKPKQIDEGADVMA 230
 170 -----KAKQKLRDNDPKQKLPVKKREYKNGKSTLDVTKMFKQIDEGADVMA 221
 221 LLDVSKNSBEDPNNKAKKIKLVTLTSKASNSDNDHKYNSRNVLMTFYREISNP 290
 222 LLDVQKQNTKMFDAKQIKKQVTLTGEPTDKEN-----HNRNSVLMTFYKVSIDP 277
 221 IDISGTEBQDLKLDLKKAKKANYDGVDLQGAHKAKEIPNKKEKKPKRHHIYLF 350
 278 IEL---TTKQVDAKKEKRVDAKQKQMDGVDLQGAHKAKEIPNKKEKKPKRHHIYLF 332
 351 SGGSTSFSTFIDHNRK-SDSKILKTRVNMNTTSNPLFPMLPIRHHNRKADMDVKKILK 391
 333 SGGSTSFSTFIDHNRK-SDSKILKTRVNMNTTSNPLFPMLPIRHHNRKADMDVKKILK 391
 411 LGQTLGAGLDNIQSTLSLSTGSSLAGAFGGGSLTEYLTLKRYSGDLKENQDVTYK 470
 392 WGRKLGIBGLNDLMDTLKLAGAASGVGGFLGGGSLTEYLTLKRYSGDLKENQDVTYK 451
 471 VGBGTHFHSFSERKKTGZLPFKSEIEPKKELPNNKNNQDSKTEWIFDKLSLTEROK 530
 452 VGBGYTHFSFSERKKTGZLPFKSEIEPKKELPNNKNNQDSKTEWIFDKLSLTEROK 506
 531 AKOETLMKLEYLEFYKREYHYVYNNLSALAAKMAQOEGITFFYSVDVTLKTSKKVKKQ 590
 507 ABEKALKVLDYFFYKRDYIYNNHLSALAAKMAQOEGITFFYSVDVTLKTSKKVKKQ 566
 591 VESTEDKKKKREDIEKRNKEDNYLQMSBGKDPFEDVDKAEKFOILLNVTVTET 650
 567 VSEBDDKKKKREDIEKRNKEDNYLQMSBGKDPFEDVDKAEKFOILLNVTVTET 626
 651 FEGGVANQKNSWQVSENNNSLHNSYKSTHKAASDASWMSLY-SNKESLTIWTSKQK 709
 627 FGNVAVSESGSWKTSLSGNSG--SSSRREVSYKGRDGSLSLSFGSTKESLTIWTSKQK 684
 710 EAEKNSSLTFKYLQVNOKQKLDKXKNR-----TKRDTSENKTSVTKOIISNTVANK 763
 685 QAEKNSSLTFKYLQVNOKQKLDKXKNR-----TKRDTSENKTSVTKOIISNTVANK 744
 745 INN-QEYKGNKLDVLYTKETVPVVDGEGVYV--PIPEKPLVEBMTPLYPALPNYPP 820
 745 INNGDTONNKKLEBVMYSYKFMPIPELDIEVVVQVPEKPLVEBMTPLYPALPNYPP 801
 821 DIPTPOLPKQEDLEISGGHPSVDIYEDTGTGAEGAKNGVSTQENRDPIDVITEDTOP 880
 802 --PTPOLPKQEDLEISGGHPSVDIYEDTGTGAEGAKNGVSTQENRDPIDVITEDTOP 859
 881 GMSGSDATVVEEDTAPKRPDVLVGGQSDPIDITEDTOPSVSGSNDAVVEEDTAPKRPD 940
 860 GMSGSDATVVEEDTAPKRPDVLVGGQSDPIDITEDTOPSVSGSNDAVVEEDTAPKRPD 919
 941 SLVGGQSDPIDITEDTOPSVSGSNDAVVEEDTAPKRPDVLVGGQSDPIDITEDTOPSVSGSNDAVVEEDTAPKRPD 1000
 920 ILVGGQSDPIDITEDTOPSVSGSNDAVVEEDTAPKRPDVLVGGQSDPIDITEDTOPSVSGSNDAVVEEDTAPKRPD 979
 1001 QAVVYTKAAHMLPASGDKRRASFTIATITIGAGLSKKRRDPEN 1046
 980 QAVVYTKAAHMLPASGDKRRASFTIATITIGAGLSKKRRDPEN 1025

RESULT 2
 533850
 fibronectin-binding protein - Streptococcus dysgalactiae
 C/Species: Streptococcus dysgalactiae
 C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
 C/Accession: S33850, S32632
 R.J. Lindgren, P.B.; McGavin, M.J.; Signaes, C.; Gues, B.; Gurnaidappa, S.; Hoecek, M.; Li
 R.J. Bloch, 214, 819-827, 1993
 Article: Two different genes coding for fibronectin-binding proteins from Streptococcus
 #Accession: S33850, S33850, M01D:93307299, PMID:831691

A: Molecule type: DNA
 A: Residues: 1-1091 <LIN>
 A: Cross-references: UNIPROT: Q06556, UNIPARC: UP100000892C0, EMBL: Z22150, NID: g288968, PI

Query Match 39.9%; Score 2138.5; DB 2; Length 1091;
 Best Local Similarity 42.3%; Pred. No. 1.5e-80;
 Matches 487; Conservative 184; Mismatches 310; Indels 171; Gaps 28;

1 MTNCKYLRKLSVGLVSGVTLIAPVTVQEVSASTSTSTSTSTASAGVSGTAASRTGS 60
 1 MTNCKYLRKLSVGLVSGVTLIAPVTVQEVSASTSTSTSTSTASAGVSGTAASRTGS 36
 61 GAAVTAATTAATTAATNGGPOSTPAVAEATPOFOAIAPAAATSTSSASSSDG----- 113
 37 -ASVTTESPAIQTEDDQSQALBEPTPAQTPSPVSAVAPAAAMADKKGIADAPAH 95
 114 -KAPQ--VTSSTPSTPAASSN-----GSHQBAAEETPQMEKXTVDKEN 160
 96 BPAPQASVQABAPAPAGVAAETVVGPTTEQARSRSKRAEIAIPQTEVEKLEVDKEN 155
 161 SKLNIKQDKTPTGSSVNNKEDTKLIRNRDGLADIVDTRTVKTNEBDITDVTVKPK 220
 156 SSLTVKQD-----EKDKLIKGRDNGQDIPDISHDVAVNGDGTMDVTLTVKPK 204
 221 QIDEGADVALLDVSKNSBEDPNNKAKKIKLVTLTSKASNSDNDHKYNSRNVLM 280
 205 QIDEGADVALLDVSKNSBEDPNNKAKKIKLVTLTSKASNSDNDHKYNSRNVLM 263
 281 MTFYREISNPIDISGTEBQDLKLDLKKAKKANYDGVDLQGAHKAKEIPNKKEKK 339
 264 IDFYKRVESSTDLGQMDAKKIDKELNRYKAKADYNGMGVDLQGAHKAKEIPNLDKEX 323
 340 KFKGRHIVLPSQGGSTFSEYELQNSVREDETKLIRLGAVTSNPLFPWPIFNHKNKI 399
 324 RSGKQKQVLSQGGSTFSEYELQNSVREDETKLIRLGAVTSNPLFPWPIFNHKNKI 380
 400 DMLDQVNVLYKQGLTLAGLDNIQSTLSLSTGSSLAGAFGGGSLTEYLTLKRYSGD 459
 381 NVVNDKQKIDPLNKLIGSGFAGAVDVAVQNTLLIGSFPFGKPNPLDYLSLDETSK 440
 460 LKQNPDTYKAVGSEHNSFSERKKTGZLPFKSEIEPKKELPNNKNNQDSKTEWIF 519
 441 INSEKFDYSRRGSEYNNRSYFDR-EVDKVGPKKLVKIK--GNLKKPQKQDVTML- 495
 520 DGLSTLTEROKKQSTLMKLEYLEFYKREYHYVYNNLSALAAKMAQOEGITFFYSVD 579
 496 SSLGANSIKKIQDMMDIKALDNLFTYRQYGFYNNHLSAQAARABREGIFVAVDT 555
 580 LKTSKRVKQVES-----TEDKKKKREDIEKRNKEDNYLQMSBGKDPFEDVDKAE 635
 556 PERAKNEINSQKSEAVYVHLKCKAEARELAKKNEKFDKTKEMSEBQK--FFQVDBP 614
 636 EKFKQILTNVYTTETPEDGVNFKNSWQVSENNNSLHNSYKSTHKAASDASWMSLY- 693
 615 EKFKQILTNVYTTETPEDGVNFKNSWQVSENNNSLHNSYKSTHKAASDASWMSLY- 664
 694 --SNKESLTIWTSKQKLDLKKAKKANYDGVDLQGAHKAKEIPNKKEKKPKRHHIYLF 750
 665 SSTSTESITWTSKQKLDLKKAKKANYDGVDLQGAHKAKEIPNKKEKKPKRHHIYLF 724
 751 VTMDIISNTVANKLNQEVKANKLDVLYTKETVPVVDGEGVYVPIPEKPLVEBMTPL 810
 725 VTEKVTSDVKYKINDKVKQKELDVSLYSEKTEPRKQVBPV--PDTPQEKRLTLL 781
 811 YPALPNYPP--TPDTP-----TPOLPKQEDLEISGGHPSVDIYEDTGTGAEG-- 855
 782 ABEKALKVLDYFFYKRDYIYNNHLSALAAKMAQOEGITFFYSVDVTLKTSKKVKKQ 840
 856 ---GAQNGVSTQENRDPIDVITEDTOPSVSGSNDAVVEEDTAPKRPDVLV 904
 841 PGSGNTVVEDQTSQEDIVLGGQGVIDFTEDSGQSGMNSHITTEDSKPSQEDVIL 900
 905 GQSGSDPIDITEDTOPSVSGSN--DATVVEEDTAPK----- 937

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2006, 01:28:04 ; Search time 316 Seconds

(without alignments)
3061.919 Million cell updates/sec

Title: US-10-771-931-1

Perfect score: 5357
Sequence: 1 MTNCKYKRLKSLVGLVSVGT.....PTIIGAGLSKRRDTBGN 1046

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Uniprot 7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5357	100.0	1046	2	084941 STREPTOCOCC
2	3575.5	66.7	1027	2	093751 STREPTOCOCC
3	3501	65.4	1026	2	048R34 STREPTOCOCC
4	3412.5	63.7	1023	2	093753 STREPTOCOCC
5	3371.5	62.9	1025	2	094507 STREPTOCOCC
6	3370	62.9	1026	2	09X386 STREPTOCOCC
7	3357.5	62.7	1025	2	P72532 STREPTOCOCC
8	3164.5	59.1	1013	2	09L953 STREPTOCOCC
9	3132.5	58.5	1015	2	093750 STREPTOCOCC
10	2973.5	55.5	1013	2	093752 STREPTOCOCC
11	2878.5	53.7	954	2	06YB06 STREPTOCOCC
12	2834	52.9	1022	2	093754 STREPTOCOCC
13	2779.5	51.9	963	2	093749 STREPTOCOCC
14	2654.5	49.6	1029	2	09X2V2 STREPTOCOCC
15	2571	48.0	862	2	09R2P2 STREPTOCOCC
16	2531	46.4	923	2	09S4J9 STREPTOCOCC
17	2488	46.4	865	2	09S4J9 STREPTOCOCC
18	2465	46.0	873	2	09S4K0 STREPTOCOCC
19	2460.5	45.9	872	2	09S4J3 STREPTOCOCC
20	2439.5	45.5	854	2	09S3P8 STREPTOCOCC
21	2329	43.5	881	2	09R3J4 STREPTOCOCC
22	2138.5	39.9	1091	2	006556 STREPTOCOCC
23	2102	39.2	746	2	08V895 STREPTOCOCC
24	1458.5	27.2	409	2	09X3B9 STREPTOCOCC
25	1391.5	26.0	456	2	09X3B9 STREPTOCOCC
26	1388.5	25.9	471	2	09X3B9 STREPTOCOCC
27	1376.5	25.7	447	2	09R2P8 STREPTOCOCC
28	1343	25.1	432	2	09R2P8 STREPTOCOCC
29	1331.5	24.9	459	2	09X3K2 STREPTOCOCC
30	1276	23.8	454	2	09S4I8 STREPTOCOCC
31	1262	23.6	423	2	09S4F6 STREPTOCOCC

ALIGNMENTS

32	1247.5	23.3	435	2	09X9C0 STREPTOCOCC
33	1238.5	23.1	437	2	09E2P7 STREPTOCOCC
34	1231.5	23.0	441	2	09E2P9 STREPTOCOCC
35	1229	22.9	410	2	09FD09 STREPTOCOCC
36	1219	22.8	253	2	09S492 STREPTOCOCC
37	1191	22.2	434	2	09X6H2 STREPTOCOCC
38	1190.5	22.2	430	2	09X6I3 STREPTOCOCC
39	1093.5	20.4	338	2	03ZJF8 STREPTOCOCC
40	1064	19.9	421	2	09E2P5 STREPTOCOCC
41	1025.5	19.1	383	2	09LSD7 STREPTOCOCC
42	845.5	15.8	564	2	0300K5 STREPTOCOCC
43	806.5	15.1	261	2	09R2U5 STREPTOCOCC
44	774.5	14.5	258	2	09X6H5 STREPTOCOCC
45	708.5	13.2	250	2	09X6H0 STREPTOCOCC

RESULT 1

084941 STREPTOCOCC PRELIMINARY; PRT; 1046 AA.

AC 084941
DT 01-NOV-1998, Integrated into UniprotKB/TREMBL.
DT 07-FEB-2006, entry version 25.
DE Serum opacity factor.
GN Name:sof2;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN:12MR.
RX MEDLINE=9922506; PubMed=10216862;
RA Courtney H.S., Hasty D.L., Li Y., Chiang H.C., Thacker J.L.,
RA Dale J.B.;
RT "Serum opacity factor is a major fibronectin-binding protein and a
RT virulence determinant of M type 2 Streptococcus pyogenes";
RL Mol. Microbiol. 32:89-98(1999).

CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license

DR EMBL; AF019890; AAC32596.1; -; Genomic_DNA.

DR GO; GO:0009886; Cell surface; IEA.
DR GO; GO:0009275; Cell wall (sensu Gram-positive Bacteria); IEA.
DR GO; GO:0016020; Membrane; IEA.

DR InterPro; IPR001899; Fibron repeat_bd.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF02986; Pn_bind; 3.
DR Pfam; PF00092; VWA; 1.
DR SMART; SM00327; VWA; 1.

DR TIGRfam; TIGR01168; Y5IRK signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
DR PROSITE; PS50234; VMPA; 1.

DR PROSITE; PS50234; VMPA; 1.
DR PROSITE; PS50234; VMPA; 1.

DR PROSITE; PS50234; VMPA; 1.
DR PROSITE; PS50234; VMPA; 1.

DR PROSITE; PS50234; VMPA; 1.
DR PROSITE; PS50234; VMPA; 1.

DR PROSITE; PS50234; VMPA; 1.
DR PROSITE; PS50234; VMPA; 1.

DR PROSITE; PS50234; VMPA; 1.
DR PROSITE; PS50234; VMPA; 1.

DR PROSITE; PS50234; VMPA; 1.
DR PROSITE; PS50234; VMPA; 1.


```

Db 61 GAAVTAATTTATNGGPOSTPAVAEATPOPOAIAVPAATSTSSASSSSSGKAPAVT 120
Qy 121 SSTSPSTPAASNGSNOEASAEETEPOTMEVEKTYVDKENSANI KQKPKPTSSVNN 180
Db 121 SSTSPSTPAASNGSNOEASAEETEPOTMEVEKTYVDKENSANI KQKPKPTSSVNN 180
Qy 181 KDTKLINRDKLADIVDTKTYVNEBGTIDVTYVVKQIDEGADVMALLDVSKTMS 240
Db 181 KDTKLINRDKLADIVDTKTYVNEBGTIDVTYVVKQIDEGADVMALLDVSKTMS 240
Qy 241 DDPNNAKTKIKVYKTLTSKASNDDEKTKNSRNSVRLMTPTFRISNPIDISGKTEBQ 300
Db 241 DDPNNAKTKIKVYKTLTSKASNDDEKTKNSRNSVRLMTPTFRISNPIDISGKTEBQ 300
Qy 301 LDKLIDLRKKAANYMGVDLOGAHKAAREIPKKEKKRKHVLPFSQGESSTYS 360
Db 301 LDKLIDLRKKAANYMGVDLOGAHKAAREIPKKEKKRKHVLPFSQGESSTYS 360
Qy 361 LONSREDEKTKLSLSGAVTSSNPPLPMPPIFNTHKNIDMLDVKNLVKGOTLAGL 420
Db 361 LONSREDEKTKLSLSGAVTSSNPPLPMPPIFNTHKNIDMLDVKNLVKGOTLAGL 420
Qy 421 DNOSTLSLSTGSSLAGAPLGGSLTEYLTLKESGDLKENQFOTYKVGSGYHNSF 480
Db 421 DNOSTLSLSTGSSLAGAPLGGSLTEYLTLKESGDLKENQFOTYKVGSGYHNSF 480
Qy 481 SERKGTGEIPKSEIBPKIKELFENNKNQNDKSWTEWIPDLSTLERIQAKOETLMKL 540
Db 481 SERKGTGEIPKSEIBPKIKELFENNKNQNDKSWTEWIPDLSTLERIQAKOETLMKL 540
Qy 541 EYLFYKRYHYNNNSALBAKAAQOEGITFISVDVTLKTSKRYKROVESTEDKKE 600
Db 541 EYLFYKRYHYNNNSALBAKAAQOEGITFISVDVTLKTSKRYKROVESTEDKKE 600
Qy 601 KQBEDIEKENREKPDNLKQMSBGQPFEDVDKAEKQILTNVYTFEFEGVANKDN 660
Db 601 KQBEDIEKENREKPDNLKQMSBGQPFEDVDKAEKQILTNVYTFEFEGVANKDN 660
Qy 661 SMOVSSENNNSLHNSYKSVTHKASDASMSVLSNKSLSWTLSKEQLKEAFENKSLTF 720
Db 661 SMOVSSENNNSLHNSYKSVTHKASDASMSVLSNKSLSWTLSKEQLKEAFENKSLTF 720
Qy 721 KYTLQVNNKQKLDLKNKRRTRDSTENKTSVTQDIISNTVYKLNQEVKGNKLDVKT 780
Db 721 KYTLQVNNKQKLDLKNKRRTRDSTENKTSVTQDIISNTVYKLNQEVKGNKLDVKT 780
Qy 781 YTKETVPVPEGEVVPPEKPLVEPMTPLYPALPNPTPTDIPPOKQBEDIEISGSG 840
Db 781 YTKETVPVPEGEVVPPEKPLVEPMTPLYPALPNPTPTDIPPOKQBEDIEISGSG 840
Qy 841 PSVDIVEDTGTAGAGANGVSTOENRBDIVDITEDTQMGSGSNDATVVEEDTAPKR 900
Db 841 PSVDIVEDTGTAGAGANGVSTOENRBDIVDITEDTQMGSGSNDATVVEEDTAPKR 900
Qy 901 DVVVGOSDPIDITEDTQPSVSGSNDATVVEEDTQPRPSLVGSGSDPIDITEDTQPG 960
Db 901 DVVVGOSDPIDITEDTQPSVSGSNDATVVEEDTQPRPSLVGSGSDPIDITEDTQPG 960
Qy 961 SSGSNGATVVEEDTQPRKVFHNEPQAPKRNQPSLSLQAPVYKAHHLPASGDKRA 1020
Db 961 SSGSNGATVVEEDTQPRKVFHNEPQAPKRNQPSLSLQAPVYKAHHLPASGDKRA 1020
Qy 1021 SFTIAPPTIIGAAGLSKKRRDTEGN 1046
Db 1021 SFTIAPPTIIGAAGLSKKRRDTEGN 1046

```

RESULT 2
03751 STRPY
ID 03751 STRPY PRELIMINARY; PRT; 1027 AA.
Integrated into UniProtKB/TrEMBL.

06:13:33 2006

```

DT 07-FEB-2006, entry version 16.
DE Serum opacity factor VT21.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21642587; PubMed=11782509;
RA Gillen C.M., Towse R.J., McMillan D.J., Delvecchio A.,
RA Stiprakaoh K.S., Currie B., Kreikemeyer B., Chatwal G.S.,
RA Walker M.J.
RT Immunological response mounted by Aboriginal Australians living in
RT the Northern Territory of Australia against Streptococcus pyogenes
RT serum opacity factor.
RL Microbiology 148:169-178(2002).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
EMBL: AB367014; AK52968.1; -, Genomic DNA.
DR GO; GO:0009886; C:cell surface; IEA.
DR GO; GO:0009275; C:cell wall (benu Gram-positive Bacteria); IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR004237; Pfam repeat_bd.
DR InterPro; IPR005877; Gpox YGIRK.
DR InterPro; IPR001899; Gram_pos_YGIRK.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF02986; Fv_bind; 3.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR SMART; SM00327; VWA; 1.
DR TIGRfam; TIGR01168; YGIRK signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
DR PROSITE; PS50234; VWF_A; 1.
SQ SEQUENCE 1027 AA; 112542 MW; 048239B8598A4C7A CRC64;

```

```

Query Match 66.7%; Score 3575.5; DB 2; Length 1027;
Best Local Similarity 70.1%; Pred. No. 2.8e-134;
Matches 751; Conservative 98; Mismatches 153; Indels 69; Gaps 26;
Qy 1 MNCKYKRLKRLSVGLVSGTLMIAPTLVQVVSAST--TSTETSTASAGVSTGTA 57
Db 1 MNCKYKRLKRLSVGLVSGTLMIAPTLVQVVSAST--TSTETSTASAGVSTGTA 57
Qy 58 TSGGAATTAATTT-----ATNGGPOSTPAVAEATPOPOAIAVPAAT 102
Db 58 TSGGAATTAATTT-----ATNGGPOSTPAVAEATPOPOAIAVPAAT 102
Qy 103 STSSASSSDGAPQAVTSTSPSTPAASNGSNOEASAEETEPOTMEVEKTYVDKENS 162
Db 103 STSSASSSDGAPQAVTSTSPSTPAASNGSNOEASAEETEPOTMEVEKTYVDKENS 162
Qy 112 STSS-SSSSDQGIPTK-TAITSPTPAASNN-SNOEASAEETEPOTMEVEKTYVDK 166
Db 112 STSS-SSSSDQGIPTK-TAITSPTPAASNN-SNOEASAEETEPOTMEVEKTYVDK 166
Qy 163 LNIKQGTERTSSVNNKQTKLIRNDRKLDIVDTKTYVNEBGTIDVTYVVKQI 222
Db 163 LNIKQGTERTSSVNNKQTKLIRNDRKLDIVDTKTYVNEBGTIDVTYVVKQI 222
Qy 223 DEGADVMALLDVSKKMSBEDPNNNAKTKIKVYKTLTSKASNDDEKTKNSRNSVRLMT 282
Db 223 DEGADVMALLDVSKKMSBEDPNNNAKTKIKVYKTLTSKASNDDEKTKNSRNSVRLMT 282
Qy 283 FTRISNPIDISGKTEBQDLKLDLRKKAANYMGVDLOGAHKAAREIPKKEKKRKH 342
Db 283 FTRISNPIDISGKTEBQDLKLDLRKKAANYMGVDLOGAHKAAREIPKKEKKRKH 342
Qy 343 KKHIVLTSQGSTSTSYELONSVRDEKTKLSLSGAVTSSNPPLPMPPIFNTHKNID 402
Db 343 KKHIVLTSQGSTSTSYELONSVRDEKTKLSLSGAVTSSNPPLPMPPIFNTHKNID 402
Qy 403 DDKKLTQTLTFTVTSNPPLPMPPIFNTHKNID 462
Db 403 DDKKLTQTLTFTVTSNPPLPMPPIFNTHKNID 462
Qy 462 DDKKLTQTLTFTVTSNPPLPMPPIFNTHKNID 462
Db 462 DDKKLTQTLTFTVTSNPPLPMPPIFNTHKNID 462

```


**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.
